

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(I) APPLICANT: Vreeland, Valerie, Ng, Kwan

(ii) TITLE OF INVENTION: Recombinant Vanadium Haloperoxidases
And Their Uses

(iii) NUMBER OF SEQUENCES: 2

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Townsend and Townsend and Crew LLP
(B) STREET: Two Embarcadero Center, Eighth Floor
(C) CITY: San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94111-3834

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US Not yet assigned
(B) FILING DATE: 08-AUG-1998
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME:
(B) REGISTRATION NUMBER:
(C) REFERENCE/DOCKET NUMBER: 023070-087100US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (415) 576-0200
(B) TELEFAX: (415) 576-0300

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2931 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 228..2258

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

	CGCGGACAAG CCTTGAAGA GAGGTTGCC AATTCAACAG AGCGAGGCC GTGAAGGTGT	60
5	GGAGGACACG TGCTACAAGC TGATCCACGA GAGCCTCAAC TTCCCTACTG ATACGGGAGT	120
	TTGTACTGCG CCGCGTTGCC AAAAACCGCA ACTTTAAACA GCGCTCGCGA GCGCCACATG	180

[illegible]

	CTTCCCACGC	ATCCACAAAA	TCGACAGTGG	TATCGCTGAG	CTTGAAT	ATG	CTT	TGC	236
						Met	Leu	Cys	
						1			
5	CAT GCA GCG GAC ACG ACA AGA GGC TCT CCT ATG CCT GAC ACC GGA GTG	284							
	His Ala Ala Asp Thr Thr Arg Gly Ser Pro Met Pro Asp Thr Gly Val								
	5 10 15								
10	CTT CGG TTG CTC ACA TCA GAG CAG CGC GCT AAA GGT TGG AGA CGC CAG	332							
	Leu Arg Leu Leu Thr Ser Glu Gln Arg Ala Lys Gly Trp Arg Arg Gln								
	20 25 30 35								
15	TTA GAG GGG GAG AAA TCA CTA GGT TTT CAT CCA AGC GAG ACG CCT TAT	380							
	Leu Glu Gly Glu Lys Ser Leu Gly Phe His Pro Ser Glu Thr Pro Tyr								
	40 45 50								
20	ATC AAG TAC TTG GAA GGC TCT GAG ACT TGG AAG AAG GTT AAG CTT CCA	428							
	Ile Lys Tyr Leu Glu Gly Ser Glu Thr Trp Lys Lys Val Lys Leu Pro								
	55 60 65								
25	ACG GAC GGC ATA TCG GCT TCC AAG ATC CTG GGT AAA ATT ATG GCC AGG	476							
	Thr Asp Gly Ile Ser Ala Ser Lys Ile Leu Gly Lys Ile Met Ala Arg								
	70 75 80								
30	GTC CGC ATC GCT ACC GCC TTG GCT GTG GTA CTG GCC GCA CCC TGT TTG	524							
	Val Arg Ile Ala Thr Ala Leu Ala Val Val Leu Ala Ala Pro Cys Leu								
	85 90 95								
35	GCA TTC GAC GAG GTC ACA GCC AGT GGT GTT TTC CCT GAG GAA CAC AAG	572							
	Ala Phe Asp Glu Val Thr Ala Ser Gly Val Phe Pro Glu Glu His Lys								
	100 105 110 115								
40	CAC ACC GGG GAG GGA AGA CAC CTC CAG ACC TGT ACA AAC TCC GAC GAT	620							
	His Thr Gly Glu Gly Arg His Leu Gln Thr Cys Thr Asn Ser Asp Asp								
	120 125 130								
45	GCG CTG GAT CCG ACG GCG CCG AAT AGA AGG GAC AAC GTA GCT TTT GCG	668							
	Ala Leu Asp Pro Thr Ala Pro Asn Arg Arg Asp Asn Val Ala Phe Ala								
	135 140 145								
50	TCG CGG CGC GAT GCC GCC AGG CGA GAA CGT GAC GGG ACA GGG ACT GTC	716							
	Ser Arg Arg Asp Ala Ala Arg Arg Glu Arg Asp Gly Thr Gly Thr Val								
	150 155 160								
55	TGC CAA ATC ACT AAC GGA GAA ACT GAT TTG GCT ACC ATG TTC CAC AAG	764							
	Cys Gln Ile Thr Asn Gly Glu Thr Asp Leu Ala Thr Met Phe His Lys								
	165 170 175								
60	TCT CTG CCA CAC GAT GAA CTG GGA CAG GTA ACC GCA GAC GAC TTC GCT	812							
	Ser Leu Pro His Asp Glu Leu Gly Gln Val Thr Ala Asp Asp Phe Ala								
	180 185 190 195								
65	ATC CTC GAG GAC TGC ATC TTA AAC GGA GAT TTC AGC ATT TGC GAG GAC	860							
	Ile Leu Glu Asp Cys Ile Leu Asn Gly Asp Phe Ser Ile Cys Glu Asp								
	200 205 210								
70	GTG CCT GCG GGA GAC CCG GCG GGT CGC CTC GTC AAT CCT ACC GCT GCG	908							

	Val	Pro	Ala	Gly	Asp	Pro	Ala	Gly	Arg	Leu	Val	Asn	Pro	Thr	Ala	Ala	
				215					220					225			
5	TTT	GCC	ATC	GAC	ATA	TCC	GGT	CCC	GCA	TTC	TCG	GCT	ACG	ACA	ATA	CCC	956
	Phe	Ala	Ile	Asp	Ile	Ser	Gly	Pro	Ala	Phe	Ser	Ala	Thr	Thr	Ile	Pro	
			230					235					240				
10	CCG	GTA	CCT	ACC	CTT	TCC	TCT	CCT	GAG	CTC	GCC	GCT	CAG	TTG	GCG	GAG	1004
	Pro	Val	Pro	Thr	Leu	Ser	Ser	Pro	Glu	Leu	Ala	Ala	Gln	Leu	Ala	Glu	
		245					250					255					
15	CTA	TAC	TGG	ATG	GCG	CTG	GCC	AGG	GAT	GTA	CCC	TTT	ATG	CAG	TAT	GGC	1052
	Leu	Tyr	Trp	Met	Ala	Leu	Ala	Arg	Asp	Val	Pro	Phe	Met	Gln	Tyr	Gly	
	260					265				270						275	
20	ACC	GAC	GAA	ATT	ACC	ACT	ACC	GCG	GCA	GCC	AAC	CTC	GCT	GGA	ATG	GGA	1100
	Thr	Asp	Glu	Ile	Thr	Thr	Thr	Ala	Ala	Ala	Asn	Leu	Ala	Gly	Met	Gly	
					280					285					290		
25	GGC	TTC	CCA	AAT	CTG	GAC	GCC	GTG	TCG	ATA	GGG	TCC	GAT	GGT	ACG	GTG	1148
	Gly	Phe	Pro	Asn	Leu	Asp	Ala	Val	Ser	Ile	Gly	Ser	Asp	Gly	Thr	Val	
				295				300						305			
30	GAC	CCG	TTC	TCC	CAG	CTC	TTC	CGA	GCG	ACC	TTC	GTT	GGT	GTT	GAA	ACG	1196
	Asp	Pro	Phe	Ser	Gln	Leu	Phe	Arg	Ala	Thr	Phe	Val	Gly	Val	Glu	Thr	
			310					315					320				
35	GGG	CCC	TTT	GTC	TCT	CAG	CTG	CTC	GTG	AAC	AGC	TTC	ACC	ATC	GAC	GCT	1244
	Gly	Pro	Phe	Val	Ser	Gln	Leu	Leu	Val	Asn	Ser	Phe	Thr	Ile	Asp	Ala	
		325					330					335					
40	ATT	ACG	GTC	GAA	CCG	AAG	CAG	GAG	ACA	TTC	GCC	CCC	GAC	TTG	AAC	TAT	1292
	Ile	Thr	Val	Glu	Pro	Lys	Gln	Glu	Thr	Phe	Ala	Pro	Asp	Leu	Asn	Tyr	
	340					345					350					355	
45	ATG	GTC	GAT	TTT	GAC	GAA	TGG	CTG	AAC	ATT	CAG	AAT	GGT	GGA	CCC	CCG	1340
	Met	Val	Asp	Phe	Asp	Glu	Trp	Leu	Asn	Ile	Gln	Asn	Gly	Gly	Pro	Pro	
					360					365					370		
50	GCC	GGC	CCC	GAA	GAG	TTA	GAC	GAA	GAG	CTG	CGT	TTT	ATC	CGT	AAC	GCC	1388
	Ala	Gly	Pro	Glu	Glu	Leu	Asp	Glu	Glu	Leu	Arg	Phe	Ile	Arg	Asn	Ala	
				375				380						385			
55	CGC	GAC	CTG	GCC	AGG	GTC	TCC	TTC	GTG	GAC	AAT	ATC	AAC	ACC	GAA	GCT	1436
	Arg	Asp	Leu	Ala	Arg	Val	Ser	Phe	Val	Asp	Asn	Ile	Asn	Thr	Glu	Ala	
			390					395					400				
60	TAT	CGC	GGG	TCT	CTT	ATC	CTA	CTT	GAG	CTG	GGA	GCC	TTC	AGC	AGG	CCC	1484
	Tyr	Arg	Gly	Ser	Leu	Ile	Leu	Leu	Glu	Leu	Gly	Ala	Phe	Ser	Arg	Pro	
		405					410					415					
65	GGT	ATC	AAC	GGT	CCA	TTC	ATC	GAC	AGT	GAT	CGG	CAG	GCG	GGC	TTC	GTC	1532
	Gly	Ile	Asn	Gly	Pro	Phe	Ile	Asp	Ser	Asp	Arg	Gln	Ala	Gly	Phe	Val	
	420					425					430				435		
70	AAC	TTC	GGC	ACG	TCT	CAC	TAC	TTC	AGA	TTG	ATA	GGT	GCC	GCC	GAG	CTG	1580
	Asn	Phe	Gly	Thr	Ser	His	Tyr	Phe	Arg	Leu	Ile	Gly	Ala	Ala	Glu	Leu	

	440	445	450	
5	GCG CAG CGT GCC TCG TGT TAC CAA AAG TGG CAG GTG CAT GGA TTT GCA Ala Gln Arg Ala Ser Cys Tyr Gln Lys Trp Gln Val His Arg Phe Ala 455 460 465	1628		
10	CGC CCC GAG GCT CTC GGG GGT ACC CTC CAC AAC ACC ATC GCG GGG GAT Arg Pro Glu Ala Leu Gly Gly Thr Leu His Asn Thr Ile Ala Gly Asp 470 475 480	1676		
15	CTA GAT GCA GAC TTC GAC ATC TCC CTT CTT GAA AAT GAT GAG CTC TTG Leu Asp Ala Asp Phe Asp Ile Ser Leu Leu Glu Asn Asp Glu Leu Leu 485 490 495	1724		
20	AAA CGT GTG GCG GAG ATA AAT GCG GCG CAG AAT CCC AAC AAC GAG GTC Lys Arg Val Ala Glu Ile Asn Ala Ala Gln Asn Pro Asn Asn Glu Val 500 505 510 515	1772		
25	ACC TAC CTT CTT CCA CAA GCT ATC CAA GTG GGA TCG CCA ACG CAC CCT Thr Tyr Leu Leu Pro Gln Ala Ile Gln Val Gly Ser Pro Thr His Pro 520 525 530	1820		
30	TCC TAC CCG TCC GGC CAC GCT ACC CAA AAT GGA GCA TTT GCC ACA GTT Ser Tyr Pro Ser Gly His Ala Thr Gln Asn Gly Ala Phe Ala Thr Val 535 540 545	1868		
35	CTG AAG GCC CTC ATT GGC CTA GAT CGG GGA GGT GAG TGC TTC CCT AAC Leu Lys Ala Leu Ile Gly Leu Asp Arg Gly Gly Glu Cys Phe Pro Asn 550 555 560	1916		
40	CCC GTG TTC CCA AGC GAT GAC GGC CTG GAA CTA ATC AAC TTC GAA GGG Pro Val Phe Pro Ser Asp Asp Gly Leu Glu Leu Ile Asn Phe Glu Gly 565 570 575	1964		
45	GCA TGC CTT ACA TAT GAG GGA GAG ATC AAC AAG CTC GCG GTC AAC GTC Ala Cys Leu Thr Tyr Glu Gly Glu Ile Asn Lys Leu Ala Val Asn Val 580 585 590 595	2012		
50	GCA TTT GGG AGG CAG ATG CTG GGC ATC CAC TAT CGG TTC GAC GGT ATC Ala Phe Gly Arg Gln Met Leu Gly Ile His Tyr Arg Phe Asp Gly Ile 600 605 610	2060		
55	CAA GGC CTA CTT CTC GGA GAG ACA ATC ACT GTA CGA ACA CTT CAC CAG Gln Gly Leu Leu Leu Gly Glu Thr Ile Thr Val Arg Thr Leu His Gln 615 620 625	2108		
60	GAG CTG ATG ACG TTC GCC GAG GAA GCC ACC TTT GAA TTC CGC TTA TTC Glu Leu Met Thr Phe Ala Glu Glu Ala Thr Phe Glu Phe Arg Leu Phe 630 635 640	2156		
65	ACC GGA GAG GTC ATC AAA CTT TTC CAG GAC GGG ACA TTC TCC ATC GAT Thr Gly Glu Val Ile Lys Leu Phe Gln Asp Gly Thr Phe Ser Ile Asp 645 650 655	2204		
70	GGA GAT ATG TGT TCC GGT TTG GTT TAC ACT GGC GTG GCG GAC TGC CAG Gly Asp Met Cys Ser Gly Leu Val Tyr Thr Gly Val Ala Asp Cys Gln 660 665 670 675	2252		

GCT TAGTGCAGAA AATAATAATT GTCGGATGCT TAAAATGCAC CCACGACCAA 2305
Ala

5 GTCGTCGAGT CACGTCGCCG GAGCATCCTT CAGCGAAAAA GGAGAGTAAC CTATATGCTA 2365
TAGAGGAGAA CCACGGAGTA CAATGCAGGT TCTTTTACCA TGTACATTGG ATTGCAGTAA 2425
GTGCGGTTAG AGAGGGATAC GTTAAACGTG CTGCGCTGTG TATATGATAC ATTTGTCATG 2485
10 GAAATATTAG AATGCGTTGA CTTGACTTCA CCATGAAATA CCATGATCGC GTGGTGTGCT 2545
GCTTTTACCT GTCGGAGCGG TACGTAAGAT GTGCTTTCTA CTGAGCCGTT TGTGTTTAGT 2605
15 CCATTCCGCG TGGCAGTGTA AACAAAGAGG ATGTAGTCTC GCCCTCAGTT TGGAGAGTAC 2665
CGTAGGTGGC AGGACGTATA TCTCTGGTAG CGGTCTGTTA AGAACTTCCA CAAGACCGTT 2725
TACGTTTGGT TGTTTAGTCG ATGCCTCTTC GTTACTTGAC CGATCCATTG AGAGTACCTG 2785
20 TACCAGTATG GTGTAAGACA TATTTTCTC CTGTTATGGA TCTGTAGAAC AGCTAGGTGT 2845
TGTTTTATAC ACAGGATGCT ATAAAATAGG GATGTTGATA ATGGCATCGG TACTCATGAA 2905
25 ACCGCAAAAT GCGGATAGAT ATTCCC 2931

20
25
TDEHD-640000

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 676 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Cys His Ala Ala Asp Thr Thr Arg Gly Ser Pro Met Pro Asp
 1 5 10 15
 Thr Gly Val Leu Arg Leu Leu Thr Ser Glu Gln Arg Ala Lys Gly Trp
 20 25 30
 Arg Arg Gln Leu Glu Gly Glu Lys Ser Leu Gly Phe His Pro Ser Glu
 35 40 45
 Thr Pro Tyr Ile Lys Tyr Leu Glu Gly Ser Glu Thr Trp Lys Lys Val
 50 55 60
 Lys Leu Pro Thr Asp Gly Ile Ser Ala Ser Lys Ile Leu Gly Lys Ile
 65 70 75 80
 Met Ala Arg Val Arg Ile Ala Thr Ala Leu Ala Val Val Leu Ala Ala
 85 90 95
 Pro Cys Leu Ala Phe Asp Glu Val Thr Ala Ser Gly Val Phe Pro Glu
 100 105 110
 Glu His Lys His Thr Gly Glu Gly Arg His Leu Gln Thr Cys Thr Asn
 115 120 125
 Ser Asp Asp Ala Leu Asp Pro Thr Ala Pro Asn Arg Arg Asp Asn Val
 130 135 140
 Ala Phe Ala Ser Arg Arg Asp Ala Ala Arg Arg Glu Arg Asp Gly Thr
 145 150 155 160
 Gly Thr Val Cys Gln Ile Thr Asn Gly Glu Thr Asp Leu Ala Thr Met
 165 170 175
 Phe His Lys Ser Leu Pro His Asp Glu Leu Gly Gln Val Thr Ala Asp
 180 185 190
 Asp Phe Ala Ile Leu Glu Asp Cys Ile Leu Asn Gly Asp Phe Ser Ile
 195 200 205
 Cys Glu Asp Val Pro Ala Gly Asp Pro Ala Gly Arg Leu Val Asn Pro
 210 215 220
 Thr Ala Ala Phe Ala Ile Asp Ile Ser Gly Pro Ala Phe Ser Ala Thr
 225 230 235 240
 Thr Ile Pro Pro Val Pro Thr Leu Ser Ser Pro Glu Leu Ala Ala Gln

2000-00000000

37
245 250 255
Leu Ala Glu Leu Tyr Trp Met Ala Leu Ala Arg Asp Val Pro Phe Met
260 265 270
5 Gln Tyr Gly Thr Asp Glu Ile Thr Thr Thr Ala Ala Ala Asn Leu Ala
275 280 285
10 Gly Met Gly Gly Phe Pro Asn Leu Asp Ala Val Ser Ile Gly Ser Asp
290 295 300
Gly Thr Val Asp Pro Phe Ser Gln Leu Phe Arg Ala Thr Phe Val Gly
305 310 315 320
15 Val Glu Thr Gly Pro Phe Val Ser Gln Leu Leu Val Asn Ser Phe Thr
325 330 335
Ile Asp Ala Ile Thr Val Glu Pro Lys Gln Glu Thr Phe Ala Pro Asp
340 345 350
Leu Asn Tyr Met Val Asp Phe Asp Glu Trp Leu Asn Ile Gln Asn Gly
355 360 365
25 Gly Pro Pro Ala Gly Pro Glu Glu Leu Asp Glu Glu Leu Arg Phe Ile
370 375 380
Arg Asn Ala Arg Asp Leu Ala Arg Val Ser Phe Val Asp Asn Ile Asn
385 390 395 400
30 Thr Glu Ala Tyr Arg Gly Ser Leu Ile Leu Leu Glu Leu Gly Ala Phe
405 410 415
Ser Arg Pro Gly Ile Asn Gly Pro Phe Ile Asp Ser Asp Arg Gln Ala
420 425 430
35 Gly Phe Val Asn Phe Gly Thr Ser His Tyr Phe Arg Leu Ile Gly Ala
435 440 445
40 Ala Glu Leu Ala Gln Arg Ala Ser Cys Tyr Gln Lys Trp Gln Val His
450 455 460
Arg Phe Ala Arg Pro Glu Ala Leu Gly Gly Thr Leu His Asn Thr Ile
465 470 475 480
45 Ala Gly Asp Leu Asp Ala Asp Phe Asp Ile Ser Leu Leu Glu Asn Asp
485 490 495
Glu Leu Leu Lys Arg Val Ala Glu Ile Asn Ala Ala Gln Asn Pro Asn
500 505 510
50 Asn Glu Val Thr Tyr Leu Leu Pro Gln Ala Ile Gln Val Gly Ser Pro
515 520 525
55 Thr His Pro Ser Tyr Pro Ser Gly His Ala Thr Gln Asn Gly Ala Phe
530 535 540
Ala Thr Val Leu Lys Ala Leu Ile Gly Leu Asp Arg Gly Gly Glu Cys

545 550 555 560
 Phe Pro Asn Pro Val Phe Pro Ser Asp Asp Gly Leu Glu Leu Ile Asn
 565 570 575
 5 Phe Glu Gly Ala Cys Leu Thr Tyr Glu Gly Glu Ile Asn Lys Leu Ala
 580 585 590
 10 Val Asn Val Ala Phe Gly Arg Gln Met Leu Gly Ile His Tyr Arg Phe
 595 600 605
 Asp Gly Ile Gln Gly Leu Leu Leu Gly Glu Thr Ile Thr Val Arg Thr
 610 615 620
 15 Leu His Gln Glu Leu Met Thr Phe Ala Glu Glu Ala Thr Phe Glu Phe
 625 630 635 640
 Arg Leu Phe Thr Gly Glu Val Ile Lys Leu Phe Gln Asp Gly Thr Phe
 645 650 655
 20 Ser Ile Asp Gly Asp Met Cys Ser Gly Leu Val Tyr Thr Gly Val Ala
 660 665 670
 25 Asp Cys Gln Ala
 675